AD_____

Award Number: W81XWH-11-1-0235

TITLE: The ZNF217 Breast Cancer Oncogene Amplified at 20q13: A Potential Marker for Invasiveness

PRINCIPAL INVESTIGATOR: Jeffrey Gregg, M.D.
Sheryl Krig, Ph.D.

CONTRACTING ORGANIZATION: University of California, Davis
Davis, CA 95618

REPORT DATE: March 2012

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
March 2012	Annual	15 March 2011 - 14 February 2012

4. TITLE AND SUBTITLE
The ZNF217 Breast Cancer Oncogene Amplified at 20q13: A Potential Marker for Invasiveness

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-11-1-0235

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Jeffrey P. Gregg
Sheryl R. Krig

E-Mail: jpgregg@ucdavis.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of California, Davis
Davis, CA 95618

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

ZNF217 is found amplified at the 20q13.2 locus in ~20% of breast tumors. We will explore ZNF217 over-expression as a surrogate marker for 20q13 amplification and a potential biomarker for prediction of invasive potential and metastatic disease. To examine overlap for ZNF217 amplification at 20q13 and ZNF217 overexpression, paraffin-embedded breast tissue samples have been requested and will be screened using aCGH to identify 100 samples that are 20q13-amplified. The ZNF217 copy number will be determined and ZNF217 protein levels measured in the 20q13-amplified tissues within the next 6 months. ZNF217 positivity will be correlated with poor patient prognosis including progression to IDC or development of distant metastases as well as disease-free and overall survival in our final report. To examine whether ZNF217 over-expression interferes with the maturation of mammary acini in 3D culture, we have extensively characterized the ZNF217_MCF10A stable cell line and describe the morphological findings in detail below. We anticipate characterization of 3D cell cultures by immunofluorescence and confocal studies will begin soon upon establishment of the ZNF217-inducible cell lines. Using the preliminary data from the cell cultures studies we have submitted an R01 to further investigate the downstream ZNF217 effector genes and cell signaling pathways driving the ZNF217-induced morphology

15. SUBJECT TERMS
ZNF217, invasive cancer, 20q13 amplification

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
U	U	U	UU	16	19b. TELEPHONE NUMBER *(include area code)*

TABLE OF CONTENTS

INTRODUCTION:

The zinc finger transcription factor, ZNF217, is found amplified at 20q13 in ~20% of breast tumors. The 20q13.2 locus is associated with aggressive features (high histological grade, DNA aneuploidy, high S-phase fraction, axillary nodal involvement). We will explore ZNF217 over-expression as a potential biomarker at 20q13 for prediction of invasive potential and metastatic disease. We hypothesize that breast tissue samples taken at the time of biopsy could be assayed for amplified ZNF217 using a simple, inexpensive clinical diagnostic test. Status of the ZNF217 biomarker would then be used for tumor prognosis and meaningful stratification to determine patient therapy for aggressive disease. Characterization of the ZNF217 phenotype in three dimensional cultures will support the hypothesis that ZNF217 amplified tumors contribute to aggressive breast disease and provide the mechanism for potential therapeutic intervention.

STATEMENT OF WORK AND PROGRESS TO DATE:

First milestone for (Krig_PI)**: Demonstrate elevated ZNF217 expression interferes with normal ductal formation, modeling ZNF217 amplification in mammary tissue**: The progress for milestone 1 is detailed in midterm report W81XWH-11-0234 (PI-Krig).

Objective 2. Assess the frequency of 20q13 copy number with ZNF217 over-expression in paraffin-embedded breast tumor samples. Comparative genomic hybridization (CGH) array will be performed to identify 100 tissue specimens with increased copies of the 20q13 locus and 50 negative control specimens without increased copies. The amplification of ZNF217 at the 20q13 locus will be quantitated by CGH analysis software. Immunohistochemistry will assay nuclear ZNF217 protein levels for multiple correlative analyses with the 20q13 copy number and patient follow-up data.

Task 1. IHC on our preliminary sample size of ~57 20q13-positive breast tumor tissue samples. (Timeframe: months 1-6.) Sophisticated aCGH software for data analysis allows determination of the specific DNA copy number for amplification of ZNF217 within the 20q13 locus (first ZNF217 “measure”). Our preliminary study on these 57 selected tissue specimens identified 69% ZNF217- amplified at 20q13 locus. We will also consider the ZNF217 false positive rate on breast tissue samples negative for the 20q13 amplification. The ZNF217 protein expression will be determined by immunohistochemistry and nuclear scoring for ZNF217 positivity in this cohort. If the IHC nuclear scoring is satisfactory, with high specificity and very low false positive rate, as determined by the blinded pathologist, our study will be deemed ongoing and we will use array CGH to identify a final sample size of 100 breast tissue samples amplified at 20q13. We previously optimized the nuclear specific immunohistochemistry (supervised by the collaborating pathologist Dr. Borowsky) using our ZNF217 polyclonal antibody and do not anticipate technical problems

Task 1: Progress to date: The original 57 paraffin tissue blocks used in our preliminary aCGHscreen were returned to the UC Davis Department of Pathology Archives. To track down these original tissue samples has been problematic and slow-going. In some cases an additional H&E slide was necessary for re-confirmation of tumor pathology. Currently, our paraffin tissue block requests are being processed and additional samples to complete the study have been ordered. Although we are behind schedule we are optimistic the study will be completed in a timely manner.

ZNF217 20q13 amplification status, as determined by further analysis of our aCGH data, has been summarized for our preliminary sample set. Each sample was classified according to the status of ZNF217 amplification:

negative (neg), questionable (maybe), or positive (+, ++). Remarkably, these studies revealed that 69% of the breast cancers examined in this cohort harbored ZNF217 amplification. The distribution of these values is illustrated in Figure 1a. An example of the region of amplification is shown in Figure 1b.

A.

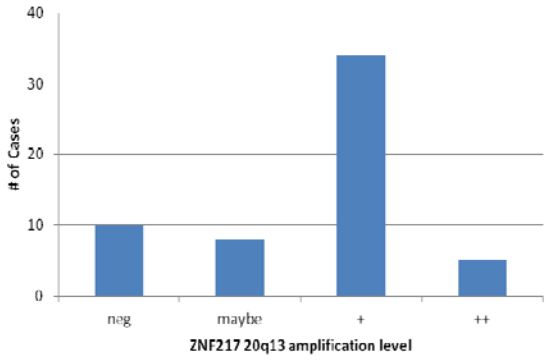
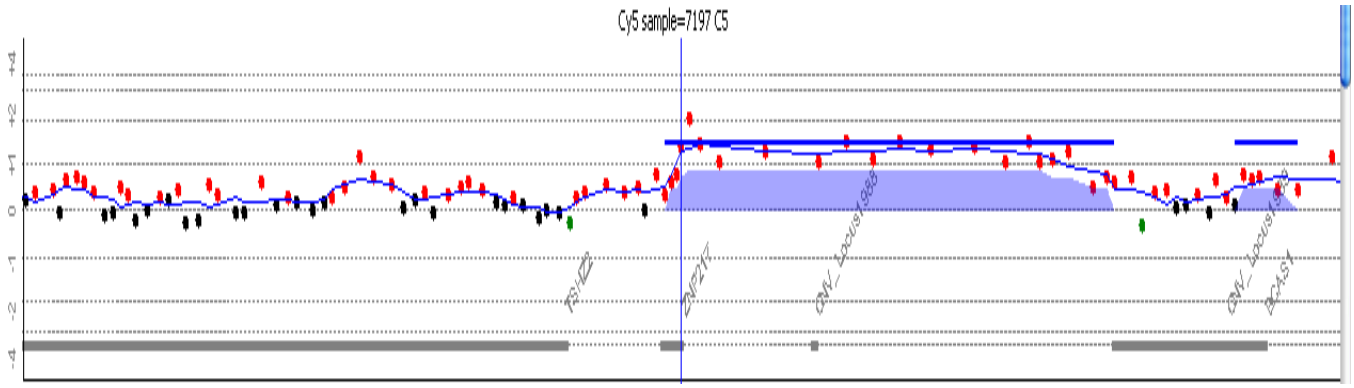


Figure 1. aCGH analysis of DNA isolated from FFPE archival tissue. A) A high proportion (69%) of the initial sample set was confirmed to be positive for the 20q13 amplification. B) Agilent CGH Analytics software was used to call regions of copy number variation. The aberration is easily detected with great precision using the Agilent 1x244K array, and the CGH Analytics software the software nicely highlights the regions of increased copy number.

B.



Task 2

Task 2: Proof of Principle will be to determine use of ZNF217 as a surrogate marker for 20q13

amplification and a biomarker for invasive disease. (Timeframe estimate: months 6-18.) An adequate number of tissue specimens with at least 10 years of clinical follow-up are readily available from our specimen repository. H&E sections will be made from each block by the UC Davis Pathology Shared Resource. The H&E section will be reviewed for each case and compared with the original diagnosis. A core of the paraffin block will be taken using a standard 6mm punch biopsy needle to a depth of approximately 2mm, DNA isolated, and aCGH performed. Based on a reported amplification rate of ~ 20%, we estimate approximately 250 specimens will be screened by aCGH to attain the final sample pool for a comprehensive study. ZNF217 copy number within the amplified 20q13 locus will be analyzed with aCGH software and IHC will measure ZNF217 protein levels for nuclei scoring in a blinded fashion by two different pathologists. Statistical assessment between the two measures of ZNF217 will determine the prognostic power for use of ZNF217 as a surrogate marker for the 20q13 amplification and marker for aggressive disease. If successful we will have established proof of principle that over-expression of ZNF217 is statistically significant for representing amplified 20q13 as a potential indicator for aggressive disease.

Although the delays with the IRB and Pathology Archive have not allowed us to collect a significant number of new cases so far, we have been able to establish and optimize a number of systems in the collection of clinical data, the isolation of high molecular-weight DNA from archival FFPE tissue, and the analysis of this DNA by aCGH.

Task 2 Progress to date:

To begin acquisition of paraffin blocks with patient follow-up data an IRB was applied for and approved August 2011 (see Appendices). Due to institutional timelines beyond our control, we were delayed in the issue of the IRB protocol. However, we now have the coded system established and the password protected database set up to ensure patient privacy on the computer system. Tissue samples have been requested from the UC Davis Department of Pathology formalin fixed paraffin embedded (FFPE) archive. The H&E section will be reviewed for each case, as necessary, and compared with the original diagnosis before we begin the DNA isolation for the aCGH assay. As the tissue samples are acquired, we will perform the aCGH for selection of the 20q13-positive samples, analyze the ZNF217 copy number, and relay the samples to the Core Facility for ZNF217 immunohistochemistry.

To streamline the identification and collection of clinical data, we created a specific list of clinical parameters to collect for all the cases in the study. This list consists of variables that are or have potential to be associated with various aspects of breast cancer, and they are subdivided into several categories: Patient Demographics, Health and Reproductive History, Tumor Data, and Treatment Data. Figure 2 shows examples of the distribution of our clinical cases across a selected set of our clinical parameters.

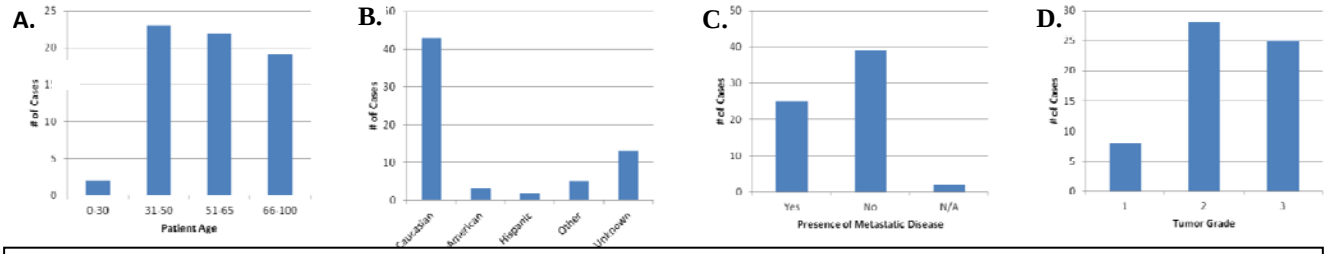


Figure 2. Clinical Parameters. A) Patient age is divided into four clinically-significant groups: 0-30, 31-50, 51-65, and 66-100 years. Distribution within our sample set is fairly even between most of the groups. B) Patient ethnicity in our current sample set is highly dominated by Caucasian women. Representation may increase for other ethnicities, but we expect that Caucasian women will continue to be in the majority. The ethnic categories shown are fairly broad, but may be refined as more cases are collected. C) Patients with non-metastatic disease represent roughly 60% of the cases. In two cases, there was no determination of lymph node status. D) This set of cases is dominated by Grade 2 and Grade 3 cancers. Grade 1 disease is represented in about 13% of the cases.

We have also optimized our sample collection process on a limited number of new samples. Our procedure now involves the use of 1.0mm – 2.0mm core biopsy punches, allowing us to obtain sample from cases that have more limiting quantities of solid tumor. High molecular-weight genomic DNA was isolated from these smaller cores using a sodium thiocyanate based isolation method. Our yields from these isolations were in the range of 4 – 6 ug of DNA per sample, making us confident that these new procedures will yield abundant DNA for the aCGH analysis.

Finally, advances in Agilent's microarray development have allowed us to move toward a higher resolution array that combines aCGH probes with probes designed to detect copy-neutral changes, such as loss of heterozygosity (LOH) and uniparental disomy (UPD), as well as single-nucleotide polymorphisms (SNPs). Instead of the previous generation's 1x244K CGH array, we will now be able to use the 2x400K CGH+SNP array. This array contains approximately 120,000 CGH probes and 60,000 SNP probes. Genotypes on this array are measured using one SNP probe per SNP, yielding approximately 5 – 10Mb resolution for LOH/UPD detection across the entire genome.

Now that a routine system of sample identification and acquisition has been established, this process will be much more streamlined and efficient. It will be reasonable to set a goal of obtaining 12 cases per week. Subsequent DNA isolation and aCGH analysis can be carried out at a similar rate. With these figures in mind, we anticipate that we will complete sample acquisition in the early part of August 2012. All aCGH arrays should be done by the start of September 2012, and analysis of the aCGH data should be completed by mid- to late September (Figure 3).

Figure 3. Revised timeline for rapid acquisition and processing of samples for aCGH. A system of regular weekly case review and archived block retrieval has been established with the Department of Pathology and will commence in mid-March. As each set of blocks is obtained and the regions of tumor are identified from the H&E slide, core punches will be obtained and DNA will be isolated. As batches of DNA are isolated, they will be placed into a weekly queue for aCGH sample preparation and array hybridization. Each batch of 12 aCGH samples will then be able to be assessed for ZNF217 20q13 amplification, and the cases will be passed on to the immunohistochemistry group. We anticipate that by early September, the 100 amplification-positive cases and 50 amplification-negative samples have been identified by aCGH. Further analysis of all the screened cases will be performed using Genomics Workbench (Agilent) during September.

Objective 3. Determine if ZNF217 has prognostic value as a biomarker for advanced disease.
(Timeframe months 18-24.)

Task 1: Clinical follow-up data garnered from our patient database will be used to correlate aggressive disease outcome with ZNF217 positivity. ZNF217 expression will be correlated with the development of local invasive disease (IDC), distant metastases, disease-free and overall survival in the combined group of 100 patients with 20q13 positive tissue samples and 50 with 20q13 negative samples. If ZNF217 has utility as a marker of invasive potential, we predict that women whose tumors score 2+ or 3+ via IHC will have an increased risk of progressing to IDC with subsequent development of distant metastases and decreased disease-free and overall survival.

Task 2: We will assess the degree to which ZNF217 status is correlated with prognostic factors such as PR, ER, and HER2 status, and has prognostic value independent of other, known factors. Additional factors from the follow-up data, such as ER, PR, HER2, and lymph node status will be analyzed for correlative significance with ZNF217 expression in collaboration with the biostatistician (see letter of collaboration with Dr. Beckett who will supervise statistical analysis). We will determine the preliminary estimate of ZNF217 correlation with specific prognostic factors. If ZNF217 correlates significantly with a subset of patients within a certain prognostic marker category, we anticipate the need for a secondary study that will determine the predictive value of ZNF217.

Progress on Objective 3: The statistical analysis of ZNF217 metrics and prognosis with patient follow-up data will be conducted upon completion of array CGH studies and IHC staining of paraffin blocks. We anticipate these analyses will begin by month 18-20.

KEY RESEARCH ACCOMPLISHMENTS:
Our cell culture studies show ZNF217 overexpression produces abnormal invasive morphology in 3D cell culture assays, establishing our hypothetical concept for further characterization using an inducible expression model.

REPORTABLE OUTCOMES:
The preliminary data generated from our ZNF217_MCF10A stable cell line formed the basis of an R01 submitted October 2011. The New Investigator grant (Krig) was reviewed and is eligible for resubmission.

CONCLUSION:

The work generated from the current DOD study will provide the support for ZNF217 as a biomarker of invasive disease. We expect to perform the statistical analysis for the correlation of ZNF217 metrics (copy number at 20q13 and protein expression by IHC) with patient follow-up data by the end of the study. The results of the study will be published.

REFERENCES

1. Debnath J, Walker SJ, Brugge JS. Akt activation disrupts mammary acinar architecture and enhances proliferation in an mTOR-dependent manner. J Cell Biol 2003; 163: 315-26.

2. Littlepage LE, Adler AS, Kouros-Mehr H, Huang G, Chou J, Krig SR, Griffith OL, Korkola JE, Qu K, Lawson DA, Xue Q, Sternlicht MD, et al. Znf217 is an oncogene that promotes a progenitor cell phenotype, metastasis and chemoresistance during breast cancer progression. Unpublished 2012.

3. Kass L, Erler JT, Dembo M, Weaver VM. Mammary epithelial cell: influence of extracellular matrix composition and organization during development and tumorigenesis. Int J Biochem Cell Biol 2007; 39: 1987-94.

4. Slorach EM, Chou J, Werb Z. Zeppo1 is a novel metastasis promoter that represses Ecadherin expression and regulates p120-catenin isoform expression and localization. Genes Dev 2011; 25: 471-84.

5. Xiang B, Muthuswamy SK. Using three-dimensional acinar structures for molecular and cell biological assays. Methods Enzymol 2006; 406: 692-701.

APPENDICES

INSTITUTIONAL REVIEW BOARD
University of California, Davis
PROTECTION OF HUMAN SUBJECTS - DECLARATION / ASSURANCE OF IRB APPROVAL

The following research study has been determined to meet the definition of human subjects research as defined by Federal Regulations and UC Davis IRB Policy and has been reviewed by the IRB in accordance with the Common Rule and any other governing regulations:

Project Title
[246349-2] The ZNF217 Breast Cancer Oncogene Amplified at 20q13: A Potential Marker for Invasiveness

Principal Investigator
Jeffrey Gregg, MD
School of Medicine

Protocol No.
246349-2

Approval Period
August 11, 2011
through August 10,
2012

Risk Level
Minimal Risk

Sponsor(s)
DOD

Status
Response/Follow-Up

Type of Review
Expedited Review

Category
5

As Principal Investigator for the above-referenced project, you assume certain responsibilities, including, but not limited to:

1. You will conduct the study according to the protocol approved by the IRB. As the PI you are ultimately responsible for the conduct of the research and the protection of rights and welfare of the human subjects. You will ensure, at all times, that you have the appropriate resources and facilities to conduct this study. You will ensure that all research personnel involved in the conduct of the study have been appropriately trained on the protection of human subjects, in addition to the study procedures.
2. Any unanticipated problems involving risks to participants or others will be reported within 5 days to the IRB or in accordance with IRB Standard Operating Procedures (SOPs).
3. Any changes in your research plan (including but not limited to advertisements) must be submitted to the IRB for review and approval prior to implementation of the change, except when necessary to eliminate immediate hazards to participants. Changes in approved research initiated without IRB approval to eliminate immediate hazards to the subject, are to be reported to the IRB in accordance with the SOP, "Reporting of Unanticipated Problems Involving Risks to Participants or Others."
4. Your protocol must be renewed prior to expiration of the study. Failure to submit renewal documents to the IRB Administration by the Administrative Due Date may result in a lapse in IRB approval or termination of the study by the IRB. All research involving human subjects must stop without on going IRB approval.
5. If you plan to collect protected health information, you are required to comply with HIPAA requirements.
6. Studies conducted at the CCRC must be reviewed and approved by the VA Research & Development Committee prior to initiation of the study. Contact the VA R&D Committee for submission requirements.
7. The UC Davis Health System requires that all investigational drugs be distributed through the UCDMC Pharmacy. You are required to provide a complete copy of the approved protocol to the Investigational Drug Service Pharmacy. A copy of the signed consent form must be submitted to the Pharmacy if investigational drugs are dispensed through the Outpatient Pharmacy.
8. For studies involving investigational drugs at Shriners Hospitals for Children Northern California, drugs must be distributed through Shriners Pharmacy. A copy of the signed consent form must be in the Pharmacy.

Name and Address of Institution
University of California, Davis
IRB Administration
CTSC Bldg, Suite 1400, Rm. 1429
2921 Stockton Blvd.
Sacramento, CA 95817

Institutional Administrator
Eric C. Mah, MHS
Director, IRB Administration
ecmah@ucdavis.edu
Phone No. (916) 703-9151
Fax No. (916) 703-9160

This Assurance, on file with the Department of Health and Human Services, covers this activity:
FWA No: 00004557

UNIVERSITY OF CALIFORNIA, DAVIS
BERKELEY • DAVIS • IRVINE • LOS ANGELES • MERCED • RIVERSIDE • SAN DIEGO • SAN FRANCISCO SANTA BARBARA • SANTA CRUZ
OFFICE OF RESEARCH
IRB Administration
TELEPHONE: 916 703-9151
FAX: 916 703-9160
SACRAMENTO, CALIFORNIA 95817

FORM W
Waiver of Research Participant's Authorization
For Use/Disclosure of PHI

The Institutional Review Board at the University of California, Davis, has approved an alteration or waiver of patient authorization to use/disclose PHI for the purpose of research. This study has been reviewed under either normal or expedited review procedures as stipulated by the HIPAA Privacy Rule. The alteration or waiver of authorization satisfies the following criteria:

• The use or disclosure of protected health care information involves no more than minimal risk to the privacy of individuals;
• Research recruitment could not practicably be conducted without the waiver or alteration; and
• The research recruitment could not practicably be conducted without access to and use of the protected health information.

A description of the protected health information, for which use or access has been determined to be necessary by the IRB, is documented in the research on file with the IRB, in accordance with the HIPAA Privacy Rule.

Human Subjects Protocol Number: 246349-2
Principal Investigator: Jeffrey Gregg, MD
Title of Study: [246349-2] The ZNF217 Breast Cancer Oncogene Amplified at 20q13: A Potential Marker for Invasiveness
IRB Approval Date: August 11, 2011

IRB Administration, Office of Research

http://www.research.ucdavis.edu/IRBAdmin

DESCRIPTION OF STUDY – Expedited Review
Version Date:07/05/11

IRB DOS – 04/26/2011
FOR IRB USE ONLY
Principal Investigator: Protocol Number (if known):
Jeffrey Gregg, MD 246349-1
Study Title:
The ZNF217 Breast Cancer Oncogene Amplified at 20q13: A Potential Marker for Invasiveness
Please address the questions below.
BACKGROUND
1. Describe the study format (e.g., pilot, phase I, phase II, multi-center, randomized, double-blinded, etc.).
This is a single-center correlative study to determine if ZNF217 is a predictor of invasiveness in breast cancer.
2. Describe the specific aims, hypothesis, and scientific problem.
Our goal is to translate a prognostic marker for stratification of tumors, assayed at the time of biopsy by ZNF217 immunohistochemistry, providing a cost-effective clinical test for measuring a new clinical marker for invasive disease. This study aims to determine ZNF217 correlation with existing prognostic factors in an effort to offer more effective tumor characterization and less toxic treatment of breast disease.

Objective 1. Model ZNF217 amplification in MCF10A mammary epithelial cells using inducible expression of ZNF217 in 3D basement membrane culture.

A. We will characterize the phenotype during the maturation of mammary acini in 3D matrigel culture overexpression of ZNF217. Using immunofluorescence and microscopy imaging techniques we will monitor apoptosis and disruption in tissue architecture.
B. Investigate whether over-expression of ZNF217 in MCF10A cells drives invasive behavior.

Objective 2. Determine the association between 20q13 copy number with ZNF217 overexpression in paraffin-embedded breast tumor samples.

A. Comparative genomic hybridization (CGH) array will be performed to identify 100 tissue specimens with extra DNA copies of the 20q13 locus and 50 specimens with no extra copies. The amplification of

ZNF217 at 20q13.3 will be quantitated by CGH analysis software.
B. We will assay both 20q13-positive and 20q13-negative (for our control group) paraffin-embedded tissue samples using immunohistochemistry (IHC) to quantify nuclear ZNF217 levels. Samples will have at least 5 years of clinical follow-up available.
C. Quantify the association between the two ZNF217 measures, aCGH and immunohistochemistry, and their correlation with 20q13.3 positivity.

Objective 3. Determine if ZNF217 has prognostic value as a biomarker for advanced disease.

A. Determine whether ZNF217 positivity is associated with poor patient prognosis including local invasion, distant metastasis, disease-free and overall survival.
B. Assess the degree to which ZNF217 expression is correlated with prognostic factors such as PR, ER, and HER2 status, and has prognostic value independent of other, known factors.

3. Provide the rationale for conducting the study and the importance of the knowledge to be gained (if not addressed in the previous question) and provide background information with references, if available.

ZNF217 is found amplified at 20q13 in ~20% of breast tumors. The 20q13.2 locus is associated with aggressive features (high histological grade, DNA aneuploidy, high S-phase fraction, axillary nodal involvement). We will explore ZNF217 over-expression as a potential biomarker at 20q13 for prediction of invasive potential and metastatic disease. We hypothesize that breast tissue samples taken at the time of biopsy could be assayed for amplified ZNF217 using a simple, inexpensive clinical diagnostic test. Status of the ZNF217 biomarker would then be used for tumor characterization and meaningful stratification to determine individual patient therapy for aggressive disease. Characterization of the ZNF217 phenotype in three dimensional culture will support the hypothesis that ZNF217 amplified tumors contribute to aggressive breast disease and provide the mechanism for potential therapeutic intervention.

4. Describe the study research methods and procedures.

Task 1. Determine ZNF217 copy number status and protein expression status in clinical breast cancer cases:

400 breast cancer tissue specimens from cases with at least 5 years of clinical follow-up will be selected from the UC Davis Department of Pathology formalin fixed paraffin embedded (FFPE) archive. Clinical data (fields for collection described in the Case Report form) will be collected from both the Electronic Medical Record (EMR) and the Laboratory Information System (LIS) for these cases and coded with internal ID numbers for anonymity. DNA will be isolated from these tissue blocks and assessed with array-CGH to determine ZNF217 copy number within the 20q13 locus. 100 ZNF217-amplified cases and 50 non-amplified cases will be selected, and ZNF217 protein expression will be analyzed in this cohort. Sections will be cut from these 150 cases and we will perform IHC and nuclear scoring for ZNF217 positivity. DNA, tissue sections, and downstream experimental data will be marked with only internal ID numbers, and once tissue sections are cut, the original tissue blocks will be returned to Pathology and the key linking the internal ID numbers to patient identifiers will be destroyed.

Task 2. Analyze clinical correlates as proof of principle for the use of ZNF217 as a surrogate marker for 20q13 amplification and a biomarker for invasive breast cancer:

Once the cases are de-linked from the patient identifiers, the clinical data fields that we selected will be analyzed with respect to the ZNF217 aCGH and protein data. Amplification/expression status will be correlated with the development of local invasive disease (IDC), distant metastases, and disease-free and overall survival.

5. Briefly describe the resources (study personnel and facilities) required for the conduct of this study and address whether they are sufficient for the conduct this study.

The array CGH and immunohistochemistry will be conducted in the UC Davis Genomics Shared Resource directed by Dr. Jeff Gregg. Dr. Gregg's laboratory has all the required equipment and resources required for the study.

DURATION & LOCATION

6. What is the duration of the study from start to end (from first subject enrollment to end of analysis of identifiable data)? How long does a given research subject participate (e.g., one study visit of 2 hours, 6 months of intervention, 5 years of intervention including follow up)?

We will collect FFPE tissue blocks from the UC Davis Pathology archive, so there will be no actual

participation time for the research subjects. The duration of the study will be three years.

7. List the specific locations for recruitment, consenting, and interventional research procedures (buildings, clinics, etc.). Note: IRB approval applies only to UC Davis organizations. Unless a written interinstitutional
agreement has been established with non-UC Davis organizations where UC Davis will serve as the IRB of Record. Contact the IRB Administration for such arrangements.

No IRB consents.

SUBJECT RECRUITMENT & CONSENTING PROCESS

8. List all the criteria for including and excluding subjects. If excluding minorities or non-English speaking subjects, also provide scientific and ethical justification for the exclusion.

Subjects to be included in the study are any women over the age of 21 with invasive breast cancer and tissue blocks in the Pathology archive. No consenting will be utilized.

9. Specify the age range of the research subjects.

Age > 21

Identify all vulnerable subject populations:
Pregnant Women (Attach Supplement) Patients as research subjects
Human Fetuses (Attach Supplement) Life-Threatening Disease
Neonates (Attach Supplement) Psychiatric patients
Prisoners Socially or Economically Disadvantaged
Children (Attach Supplement) Employees under your supervision
Cognitively impaired Students for whom you are an Instructor
Non-English speaking subjects
10.
Other (Specify):

11. Describe the additional safeguards/accommodations that will be put into place for any and all vulnerable population(s) indicated in the previous question.

12. Indicate how many subjects will be recruited and provide statistical justification for the number of subjects required.

400 subjects (FFPE blocks) will be used in this study**.** ZNF217 is found amplified at the 20q13 locus in >20% of breast tumors and is a candidate oncogene driving selection of the 20q13 amplicon. By analyzing 400 subjects, we expect to obtain 100 ZNF positive samples. Statistical analysis will then be performed on the 20q13-positive and negative samples to assess the association between ZNF217 over-expression and increased copy number in 20q13-positive and negative breast tissue samples, and to assess the relationship between ZNF217 findings and 20q13 positivity. We will calculate the proportion and 95% confidence intervals of ZNF217 positivity (CGH, IHC) separately for 20q13-positive tissue samples (true positive) and 20q13-negative samples (false positive rate). With 100 positive tissue samples, we will have a 95% CI of width at most ı0.10 for the true positivity, and at most ı0.14 for false positive based on 50 negative tissue samples. If a true positive rate of 40% is too low to pursue but a rate of 55% or higher is considered encouraging, we would have 91% power to detect such a promising level. If we consider a rate of 60% false positive to be too high to pursue but 40% to be a minimum for encouraging findings, we will have 89% power to detect a rate as low as 40% with a one-sided test at level 0.05. We will also assess the association between the two ZNF measures, across all 150 samples, by estimating overall percent agreement, kappa, and other measures of agreement. This sample size is large enough to expect very high precision for estimates of association that exceed 50-60%.

We will then determine whether ZNF217 positivity is associated with poor clinical prognosis including local invasion, distant metastasis, disease-free and overall survival, in the combined group of 100 patients with 20q13 positive tissue samples and 50 with 20q13 negative samples. For each outcome, we will consider ZNF217 positive status, separately for CGH and IHC assays, as a predictor in

regression models. The primary clinical endpoint will be progression-free survival (PFS), and association will be assessed using a proportional hazards model (two-sided test, level 0.05). Recent estimates suggest that the ten-year PFS will be no greater than 83% in those with low expression and

possibly as low as 60%. If we assume that 40% of our 100 20q13 positive samples and 70% of our 50 20q13 negative samples will have low ZNF217 expression, then we would have 75 low-expression and 75 high-expression samples for this study. With this sample size, we would have 80% power to detect a 2.3-fold increase in hazard for the high-ZNF217 group, corresponding to a drop from 83% PFS at 5 years to 65%, or 81% power to detect a decrease in PFS from 60% to 40% (1.8-fold increase in hazard). Additional analyses will use proportional hazards models to look at overall survival, development of local invasive disease, and distant metastases as endpoints.
Relationships between ZNF217 outcome and prognostic factors such as PR, ER and HER2 status will be examined by exact tests of association for categorical variables. The degree to which ZNF217 is predictive of outcome, independent of these prognostic factors, will be examined by logistic regression for dichotomous outcomes (response/nonresponse) and by Cox proportional hazards models for survival outcomes (progression-free survival). In each case, we will examine whether adding ZNF217 status to a predictive model with other important prognostic factorsadds significantly to the predictive value ($p<0.05$) beyond the information contained by other factors.
Power calculations used the oncology study design resources provided by the Southwest Oncology Group web site, http://www.swogstat.org/stat/public.
13. Provide a detailed description of the methods and process to be used for recruitment. (Attach a copy of all recruitment materials. If recruiting patients, HIPPA applies. Attach supplement.)
We will not be recruiting patients. We will use the LIS to search for and identify 400 invasive breast cancer cases where the patient is female, where the patient was over the age of 21 at the time of diagnosis, where there is at least 5 years of clinical follow-up, and there is FFPE tissue available in the archive. Once the cases are identified, they will be coded with an internal Case ID number such that samples derived from the FFPE tissue and data derived from these samples (including clinical data) are only associated with this coded Case ID number. The key that links the Case ID number to patient identifiers (EMR ID, Surgical Pathology ID) will be kept separate from all other data and will be destroyed once all downstream samples (DNA, tissue sections) are collected.
Indicate the method(s) you will employ to initially obtain the consent of subjects:
Provide the consent form in person, give prospective subject sufficient time to review the consent form and discuss with friends/family.
Consent form will be mailed, and consent process will occur over the telephone, person obtaining consent signs and dates form when s/he receives it back.
Request signed informed consent be waived by the IRB for screening purposes only. Consent will be obtained from eligible participants. (Attach Supplement)
Request signed informed consent be waived by the IRB. (Attach Supplement)
Request an alteration of informed consent. (Attach Supplement)
(Please describe):
14.
Other (Please describe and, if applicable, attach Supplement): Requesting waiver of consent (Waiver/Alteration of Informed Consent attached)
Consent Process Assurances (Please read and check the following indicating your agreement.)
At the beginning of every study visit, you will tell the subject what to expect and ensure the subject understands.
At the end of the study visit, you will explain what happens at the next visit, or when the study is over.
You will ask and ensure the subject wants to continue participation and knows s/he may refuse to participate.
15.
You will respect the rights and welfare of research subjects.

BENEFITS & RISKS OF THE RESEARCH
16. Address if there are any potential benefits to individual subjects or to the particular group or a particular community, group or society. Note: Payment for participation is not considered a benefit.
This study may benefit future breast cancer patients by potentially offering a more effective tumor characterization and more effective and/or less toxic treatment of breast disease.
Describe your plan for protecting and maintaining subject privacy and confidentiality.
The data and/or specimens will be directly labeled/recorded with the personal identifying information when acquired.
The data and/or specimens will be labeled with a code that the research team can link to personal identifying information when acquired. The code sheet will be secured and kept separate from the

dataset.
The data and/or specimens will not be labeled with any personal identifying information, nor with a code that this research team can link to personal identifying information.
17.
Other (Please describe):
18. Research Material Security: Describe your recordkeeping system and plan for securing and maintaining source documents and specimen (e.g., data management, consent forms, IND or IDE applications and approvals, sponsor notifications and monitor visit reports, DSMB reports, documentation of performance and adherence to the trial). In addition, please identify who will have access to the system.

The Laboratory Information System (LIS) in the Pathology department will be utilized to select 400 invasive breast cancer cases that are at least 5 years old. Once selected, these cases will be coded with an internal case ID number. The key to this coding system will be kept on a dedicated, secure, password-protected computer, in a password-protected Excel file. One hard (paper) copy of this key will be printed and kept in a locked cabinet in the laboratory. Clinical data (fields outlined in the Case Report form) will be obtained from the EMR and will be entered into an independent Excel spreadsheet that is kept on a separate computer and only identifies the cases using only the internal case ID number. All DNA samples, tissue sections, and downstream aCGH and IHC data will also be identified only by the internal case ID number. The tissue blocks obtained from Pathology will be marked with the surgical pathology ID number, but these will be kept in a locked cabinet during our custody of them and will be returned to Pathology as soon as the DNA is isolated and the tissue sections cut. Once the DNA is isolated and the tissue sections are cut, the key that links the internal case ID numbers with the patient identifiers will be deleted and the hard drive of the computer will be reformatted. At this time, the one hard copy of the key will also be destroyed (shredded).

Access to the computer housing the key to the ID codes will be limited to Dr. Jeffrey Gregg, Dr. Sheryl Krig, and Stephenie Liu. They will be the only people with the password to the computer and also the password for the Excel file. The computer itself will be in a locked office to which only Dr. Gregg's laboratory personnel have the key. The cabinet housing the tissue blocks and the one hard copy of the key to the ID codes will be locked at all times and will be accessible only to Dr. Gregg, Dr. Krig, and Stephenie Liu.

19. Address all the expected risks (e.g., physical, financial, social and psychological) of participating in the study.

Due to the nature of the study, we do not anticipate any adverse or serious adverse events. Serious adverse events will be reported to the IRB within 5 days of the investigator becoming aware of the event. Adverse events will be reported to the IRB at the renewal.

The risk of a breach of confidentiality will be minimized through engineering controls (dedicated computer, locked doors, limited personnel access, password-protections). In addition, no patient names, dates, or contact information will be collected. However, if a breach of confidentiality does occur, it will be reported to the IRB as an unexpected event.

COSTS AND COMPENSATION TO THE SUBJECTS

Identify study related costs to be borne by the: NONE

Subject Insurer

No costs

Blood draws (specify which blood draws):

Travel

Parking

20.

Other (specify):

21. Address how the costs to the subject/insurer are justified.

22. Will subjects be compensated for participation in the study? Yes No

If yes, describe the amount and type of compensation that will be paid to subjects and clarify how compensation will be pro-rated if the subject does not complete all study visits.

FINANCIAL INTERESTS

Disclose all financial interests for you and your key research personnel in the study, sponsor, or sponsor's competitors.

I and the co-Investigators (including all key personnel) involved with this protocol do not have a financial interest in the study, products, sponsor, and/or sponsor's competitors.

There are reportable financial conflicts of interests and all interests have been disclosed to the Conflict of Interest Committee. (Attach COIC application and/or approval)

23.
Other (Include and describe all outside income Industry sources):